

## Ethnic differences in the Goodenough–Harris draw-a-man and draw-a-woman tests

A. E. DUGDALE AND S. T. CHEN

*Department of Child Health, University of Queensland, and the Department of Paediatrics, University of Malaya*

**SUMMARY** The draw-a-man (DAM) and draw-a-woman (DAW) tests were given to 307 school-children in Petaling Jaya, Malaysia. The children were ethnically Malay, Chinese, or Indian (Tamil), and all came from lower socioeconomic groups. The standard scores of the Chinese children averaged 118 in the DAM and 112 in the DAW tests. These scores were significantly better than the American standards. Malay children scored significantly lower than Chinese, and Tamil children scored lower again. The nutritional status of the children had no influence on the scores. Chinese and Tamil children scored better in the DAM than the DAW, while in Malay boys the reverse was true. Malay children tended to emphasise clothing in the DAM, but Chinese and Tamil children scored better on items relating to facial features and body proportions. The Goodenough–Harris draw-a-person tests are obviously not culture-free, but the causes of ethnic differences have not been elucidated.

A child's performance in an intelligence test depends on his cultural background, his genetic intellectual potential, and his motivation. Most tests have been designed to match the cultural background of middle-class American or Western European children and cannot safely be used with other groups. However, one test which has the potential of being truly cross-cultural is the Goodenough–Harris draw-a-person test. The human form is universally available as a model, and figure drawing is common among children of all ethnic groups. We used this test on groups of schoolchildren from similar socioeconomic levels but very different cultural backgrounds to try to find whether significant variability existed.

### Methods

Malaysia is particularly suited for cross-cultural studies. In the cities, there are peoples of Malay, Chinese, and Indian backgrounds living side by side, but still retaining their own vernacular and customs. Within each ethnic group, the families range from affluent to very poor, and nutritional

levels also vary widely. Children from poorer families usually attend vernacular schools intermittently and leave early.

There were data on 307 children, although not all information was available for some. The data used in this study were (1) Age; this was obtained from birth certificates and ranged from 6½ to 10½ years. (2) Sex of the child. (3) Ethnic group; there was no intermarriage or ethnic mixing among the groups studied. (4) Type of school. Malay and Indian children were taught in their vernacular language, Malay and Tamil respectively. Chinese children attending Chinese language schools were taught in Mandarin, although their vernacular languages were generally Cantonese or Hokkien. (5) Number of children currently at home. (6) Family income; this was obtained by questioning the parents. (7) Number of years' schooling of each parent. (8) Height and weight of the child. This was converted to a percentage of the Stuart and Stevenson standard for a child of that age and sex (Nelson, 1964). (9) The Goodenough–Harris draw-a-man (DAM) and draw-a-woman (DAW) tests; these were done in the school classrooms, and raw scores were converted to standard scores using the tables given by Harris (1963).

The children whose drawings have been used in this study were all attending government primary

University Paediatric Unit, Mater Children's Hospital, Queensland, Australia

A. E. DUGDALE, reader in child health

Department of Paediatrics, University of Malaya

S. T. CHEN, associate professor of paediatrics

schools in Petaling Jaya, a satellite city of Kuala Lumpur, the capital of Malaysia. These schools are free and, because at the time of the study most teaching was done in vernacular language, each school had predominantly pupils from one ethnic group (for further details see Dugdale and Chen, 1977).

Medical students doing their paediatric posting in the University of Malaya ran a School Health Service under the supervision of S.T.C. Children were seen class by class by the medical students. The DAM and DAW tests were done in the classroom by the whole class.

Because the child's name appeared on his drawings, it was not possible to score the test 'blind'. However, the scoring method was objective and standardised, so knowledge of the race of the child should have had little or no effect on the scoring.

## Results

**Overall scores in the DAM test.** The mean values of the standard scores in the DAM and DAW tests are shown in Table 1. The scores are standardised on American children to have a mean of 100 and a SD of 15 points. The distribution of the DAM scores is shown in Fig. 1.

When compared with expected American standard scores, the Chinese children, boys and girls, scored higher (Kolmogorov-Smirnov 2-sample test,  $P < 0.01$ ) than American standards, but Chinese boys and girls did not differ significantly one from the other. On the other hand, Tamil children scored lower than American standards ( $P < 0.05$ ) and again boys did not differ from girls. Among Malay children, neither boys nor girls differed from one another or from American standards, although in

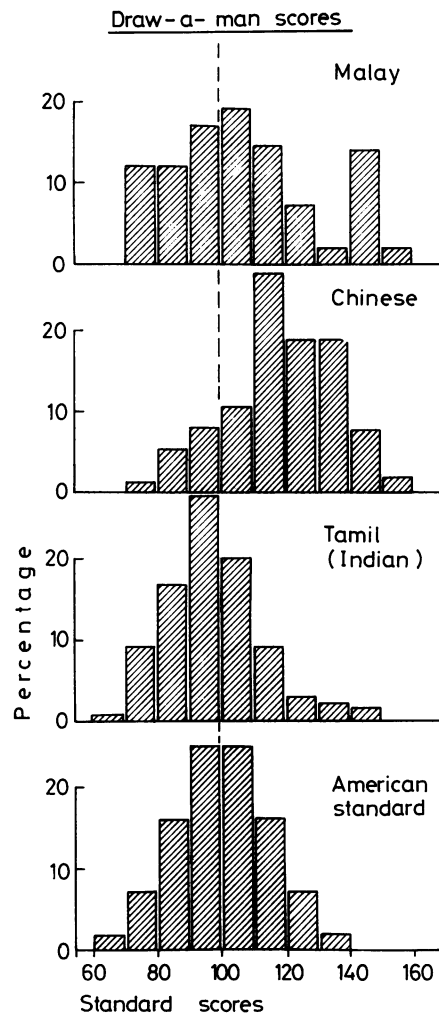


Fig. 1 Distributions of scores in the DAM test in children of different ethnic groups. American standard score is shown at the bottom.

Table 1 Numbers of children in each group who completed the DAM and DAW tests, with the mean standard scores

Group	Test			
	Draw-a-man		Draw-a-woman	
	No.	Mean score	No.	Mean score
Malay				
Boys	13	102.0	13	106.8
Girls	29	111.9	29	115.17
Chinese				
Boys	57	116.2	57	110.1
Girls	64	120.6	64	114.1
Tamil				
Boys	65	96.6	64	91.5
Girls	79	96.4	77	94.6

Malay girls the higher mean score approached the level of significance ( $0.1 > P > 0.05$ ). Within the three ethnic groups Chinese children scored higher than Malays ( $0.05 > P > 0.025$ ), who in turn scored higher than the Tamil children ( $0.05 > P > 0.025$ ). A few Tamil children drew men in long robes (see Fig. 3c) which would detract from the score. However, even allowing for this, Tamil children scored badly.

**Overall scores in DAW test.** The mean scores are shown in Table 1 and the distributions are shown in Fig. 2.

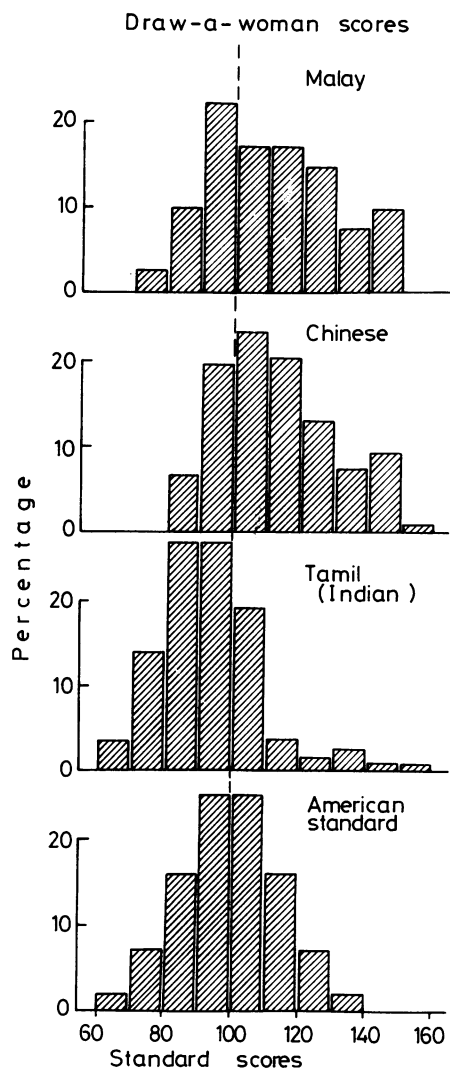


Fig. 2 Distributions of scores in the DAW test. American standard score is shown at the bottom.

Within each ethnic group, there were no significant differences between boys and girls. As with the DAM test, Chinese children scored higher than the American standards ( $P < 0.01$ ) and Tamil children lower ( $P < 0.01$ ) while Malay children tended to be better, but this did not quite reach the  $P = 0.05$  level of significance. There were no significant differences between Malay and Chinese children, but these two groups had much higher scores than Tamil children ( $P < 0.001$ ).

**Effects of nutrition and other environmental factors on scores.** Because of the high correlation ( $r = 0.80$ ) between DAM and DAW scores the mean value of the DAM and DAW was used as the dependent variable in the analysis. Multiple regression analysis was performed using the percentage present weight-for-age of the child as an indicator of nutritional status, family size, family income, and educational achievement of the parents as independent variables. These data were available on 263 children. Only 18% of the variance in the DAM and DAW scores was accounted for by these factors, and family income was the only significant single factor. Further testing with the extended median test showed that sex of the child had no effect on the scores. Although all children came from poor families, the Chinese families had slightly higher monthly incomes than the Malays, who again had slightly higher incomes than Indians. The DAM and DAW scores were recalculated to make allowance for these variations in income, but the statistically significant differences in the scores were unchanged. It is interesting that the current nutritional status of the child did not affect the score.

**Relationship between DAM and DAW scores.** An interesting pattern was found in the levels of the DAM and DAW scores among children of different ethnic groups (Table 2).

Among the Chinese children, both boys and girls scored significantly higher in the DAM than in the DAW test. A similar result was found in Tamil children. However, Malay boys scored better in the DAW than the DAM ( $P \approx 0.02$ ) and Malay girls showed no significant difference between the two tests.

**Differences in patterns of drawing.** Apart from diversity in overall scores, children from various

Table 2 Pattern of scoring in the DAM and DAW. Numbers in each group were treated using  $\chi^2$  against the hypothesis that the two patterns of scoring should be equal. Children with tied scores are omitted.

Group	No. of children with each pattern of scoring		Statistical significance P
	DAM > DAW	DAM < DAW	
Malay			
Boys	2	10	<0.05
Girls	15	13	NS
Chinese			
Boys	36	19	<0.05
Girls	47	15	<0.001
Tamil			
Boys	46	16	<0.01
Girls	53	24	<0.001

ethnic groups tended to emphasise different features of their drawings. We examined this by analysing the 73 individual items of the DAM score. For

Table 3 *Pattern of scoring of individual items in the DAM*

Group of children	Less commonly found features	More commonly found features
Malay	23 Proportion and position of ears 62 All body parts in proportion 18 Hair I 19 Hair II 22 Ears shown	55 Clothing I 56 Clothing II 58 Clothing IV 59 Clothing V 52 Tapering of arms
Chinese		15 Projection of chin 29 Wrist and ankle shown 50 Proportions of face
Indian	55 Clothing I	38 Knee shown 22 Ears shown 13 Nose and lips in two dimensions 48 Proportions of head

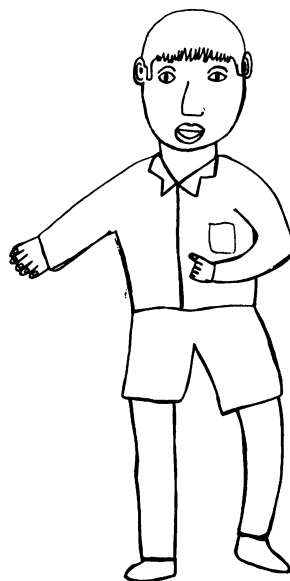
children in each ethnic group the items have been ranked in order of decreasing frequency of scoring. That is, an item which was scored positive in every child is ranked first, and items in which no child scored are ranked last. The rankings obtained for the three groups of children were compared, and items showing wide variation in ranking noted. In this way different patterns of drawing could be detected without interference of the variation in total scores. There were 7 items where the difference in rankings among the three groups was  $\geq 20$  (items 15, 18, 19, 22, 23, 38, and 52) and a further 9 where the rank differences were  $\geq 15$  (items 13, 29, 48, 50, 55, 56, 58, 59, and 62). When these differences were tabulated a distinct pattern emerged (Table 3).

The most striking differences were among the Malay children who tended to score well on clothing and poorly on body proportions and facial features. Indian (Tamil) and Chinese children tended to emphasise facial features and there may be a tendency for Tamil children to score poorly on items related to clothing. The various items in the scoring system are not independent, so statistical tests could not be used to find the significance of these patterns. Examples of drawings showing these differences are shown in Fig. 3a, b, c.

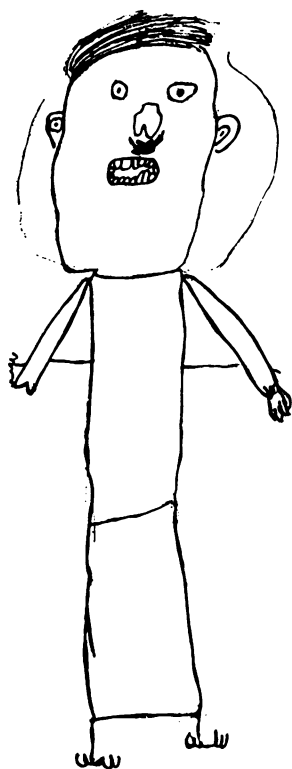


Fig. 3 Examples of drawings by three 10-year-old children of different ethnic groups.

(a) The picture of a man by a Malay child emphasises the clothing. Note also the car in the background. (IQ equivalent 105).



(b) Chinese child has drawn a man with very simple clothing but excellent facial features. (IQ equivalent 124).



(c) *Tamil child has drawn a less satisfactory man; even allowing for the long robe which is worn by some Indian men. (IQ equivalent 79).*

These drawings have been copied by an artist, as the original pencil drawings did not reproduce photographically.

## Discussion

Although drawings of the human figure appear to be a valid measure of a child's growth in conceptual complexity (Harris, 1963), there must also be some effects of cultural influences and personality characteristics. This study confirms the work of Sundberg and Ballinger (1968) and of earlier authors in showing that there are influences which invalidate the draw-a-person as a culture-free test. The standard scores for the DAM and DAW tests were lowest in the Tamil children and these differences could not be accounted for by the measured nutritional or social factors. All such measures must be incomplete and the fact that no causative factors have been identified does not mean that they do not exist. It would certainly be important to find why Chinese children from poor socioeconomic backgrounds score so much better

than American children on whom the tests were standardised.

Although the factors influencing the scores cannot be identified, there is ample evidence that they exist and differ among the three ethnic groups we studied. The evidence takes two main forms. When the DAM and DAW scores are compared in individual children, more Chinese and Tamil children scored better in the DAM than the DAW. It is possible that the Tamil children were disadvantaged in the DAW as the customary ethnic female dress, the sari, does not promote good scoring in the DAW, but this could not apply to the Chinese children. Nor does it explain why Malay girls showed no difference between DAM and DAW, but Malay boys scored higher in the DAW than the DAM.

The other feature of the DAM is the different features emphasised by the children of different ethnic groups. Malay children tended to score higher on items related to clothing, while Chinese and Tamil children were apparently more aware of facial features and body proportions. It is possible that the response of Malay children was influenced by the traditional Muslim prohibition on portraits of human subjects, even though this prohibition is by no means rigidly observed.

Any speculation about the causes of the differences in overall scores or in the patterns of response would be unwise in view of the lack of factual data. The Hopi Indians have also been shown to exceed DAM norms (Havighurst *et al.*, 1946) and Nepalese children show different patterns of response (Sundberg and Ballinger, 1968), so such variations are obviously not confined to these Malaysian children.

The observed differences in the DAM and DAW tests probably reflect variations in cultural background among the children of different races (Kuttner, 1970). Such variations are likely to affect a child's motivation and performance at school and perhaps in later life. It is currently fashionable to deny the existence of ethnic differences in intellectual skills. However, it would be more realistic to accept that different cultures may have particular strengths and weaknesses, and adjust educational methods and goals to enhance the strengths and compensate for the weaknesses.

We thank Miss Mei Li Li for scoring the drawings and Mrs Dianne Chandler for her assistance.

## References

- Dugdale, A. E., and Chen, S. T. (1977). Factors influencing school achievement of children from low socio-economic groups in Malaysia. *International Journal of Psychology*, 12, 39-49.

- Harris, D. B. (1963). *Children's Drawings as Measures of Intellectual Maturity. A Revision and Extension of the Goodenough Draw-a-Man Test*. Harcourt, Brace, and World: New York.
- Havighurst, R. J., Gunther, M. K., and Pratt, I. E. (1946). Environment and the draw-a-man test: the performance of Indian children. *Journal of Abnormal Social Psychology*, **41**, 50-63.
- Kuttner, R. E. (1970). Comparative performance of disadvantaged ethnic and racial groups (abstract). *Psychological Reports*, **27**, 372.
- Nelson, W. E., editor (1964). *Textbook of Pediatrics*, eighth edition. Saunders: Philadelphia.
- Sundberg, N., and Ballinger, T. (1968). Nepalese children's cognitive development as revealed by drawings of man, woman, and self. *Child Development*, **39**, 969-985.

Correspondence to Dr A. E. Dugdale, University Paediatric Unit, Mater Children's Hospital, Mater Hill, Queensland 4101, Australia.

Received 16 February 1979